

## OTKA pályázat záró beszámolója

Cím:

*Video történések felismerése automatikusan detektált adatbázis asszociációk segítségével*

Száma: T 49001

Téma vezetője: Szirányi Tamás

Tartam: 2005. február 1 – 2007. december 31.

### Beszámoló tartalma

OTKA pályázat záró beszámolója.....	1
Összefoglaló .....	2
Abstract .....	2
Pályázatban vállalt feladatok:.....	3
Tervezett munkaszakaszok:.....	3
A munka során elért eredmények:.....	4
1. év: .....	4
2. év: .....	4
3. év: .....	6
A munkaterv az eredmények tükrében.....	7
Személyi változások .....	7
Függelék:           Néhány fontosabb és jellemző eredmény rövid bemutatása .....	8
Fókuszterkép mérése a főszereplő szegmentálásához.....	8
Tükör és árnyék vetítési pontjának meghatározása statisztikus optimalizációval, előzetes modell nélkül.....	10
Színproblémák és metrikák .....	12
Kamerák regisztrációja akkumulált mozgás-statisztikák segítségével.....	14
Mozgó kamera képének szegmentációs problémái .....	16
OTKA támogatás feltüntetésével készült publikációk: .....	18
További kapcsolódó publikációk: .....	18

## Összefoglaló

*Video történések felismerése automatikusan detektált adatbázis asszociációk segítségével*

Elvileg is új algoritmusokat dolgoztunk ki videó részletek elemzéséhez.

Az egyik ilyen eljárás szerint a képek relatív fókuszterképét határozzuk meg előzetes modell vagy bármilyen paraméter ismerete nélkül. Ehhez egy elvileg is új hiba-mértéket vezettünk be, ami a vak-dekonvolúciós iterációk során a hibának az ortogonalitási feltételtől való eltérését mutatja.

Kidolgoztuk a mozgó és álló objektumokat elkülönítését mozgó kamera esetére, miközben azt is becsülni kell, hogy a teljes megfigyelt területből mit lát éppen a kamera.

Automatikus képillesztő eljárásokat fejlesztettünk ki, amelyek általános megoldást biztosítanak széles bázistávolságú sztereó-képek illesztésére tetszőleges körülmények esetén. A mozgás-nyomok statisztikai összehasonlításával az eseménytér szerkezetének geometriai adataira tehetünk becslést előfeltevés és modell nélkül. Új módszert adtunk az egynézetű képen levő tükrök, illetve az árnyékot adó fényforrás vetítési pontjának meghatározására. Az eljárás során statisztikai korrelációtérképekre illesztett parametrikus hipotézismodelleket optimalizálunk a vetítési hiba segítségével. A kapott eredmények jól jellemzik a színhely eseményterének elrendezését.

A videóképek mozgó alakzatainak eredményes detektáláshoz egyértelmű eredményt adó kutatást folytattunk az optimális színmetrika kialakítására.

Módszereket dolgoztunk ki videófelvételek vizuális szempontból fontos területeinek és eseményeinek automatikus elkülönítésére és a módszerek felhasználására a képi visszakeresésben.

## Abstract

*Classification of video events through automatically detected categories of video database*

We have introduced theoretically new algorithms for analysing video shots and events.

We present an automatic focus area estimation method, working with a single image without a priori information about the image, the camera, or the scene. It produces relative focus maps by localized blind deconvolution and a new residual error-based classification. Evaluation and comparison is performed and applicability is shown through image indexing.

We have developed segmentation method for arbitrary foreground objects in case of indefinitely moving cameras.

A new motion-based method is presented for automatic registration of images in multicamera systems, to permit synthesis of wide-baseline composite views. Our approach does not need any *a priori* information about the scene, the appearance of objects in the scene, or their motion. We introduce an entropy-based preselection of motion histories and an iterative Bayesian assignment of corresponding image areas. Correlated point-histories and data-set optimization lead to the matching of the different views.

An automatic method is presented using motion statistics to determine vanishing-point position for the geometrical modelling of reflective surfaces or cast shadows, even in cases where the correspondences are corrupted with considerable amounts of noise.

We proposed an optimal colour space for modelling problem of cast shadows in video sequences. The performance of shadow detection can be improved significantly through appropriate colour space selection, applied in an MRF framework.

## **Pályázatban vállalt feladatok:**

Olyan algoritmikus eljárásokat kívánunk létrehozni, amelyek a video-forrásból tulajdonságelemzés útján kapott jellemzők statisztikai becslésével jó hatásfokkal csoportosítják illetve felismerésre alkalmasan leírják a képsor eseményét. A csoportokat a megelőző tanítási folyamat során kell kialakítani a video adatbázis tartalmi besorolása (humán kategorizálás) és a gépi mérési eljárások összevetésével. A megelőző tanítást nagyszámú digitalizált videón kívánjuk végezni.

A képi történéseket előbb jellemzőkkel leírjuk, majd a legfontosabbakat kiválasztva elvégezzük a csoportosításokat. A fenti művelet sor matematikailag megfelel a Principal Component Analysis (PCA) vagy a Most Discriminant Information (MDI) szerinti felbontásnak. Ilyen módon a lényegesen összetettebb képi eseménysort parametrizáljuk ismert eseménysorok csoportosítása és az azzal történő kategorizálás útján. Az eseményeket bekövetkezésük és egymásba történő átmenetük alapján valószínűségekkkel jellemezhetjük, és a tervezett algoritmikus megoldások alapvetően a rejtett Markov-modelleken (HMM) és a Support Vector Machine (SVM) azonosítási technikákon alapulnak.

A videók jellemzéséhez szükséges paramétereket az MPEG-7 szabványában leírtakkal összhangban, XML adatbázisban adjuk meg. Az ehhez vezető eszköztár, szabványosítás és adatbázis kezelés nem része ennek a feladatnak, hanem más kapcsolódó munkáink szolgáltatják a tesztelési környezetet.

A felismerés nem csupán csoportosítást jelent, hanem az eseménysorban történtek jellegére is becslést kell tenni. Ehhez szükség van a vizuális-, hang- és esetleg egyéb (forrás jellege, előzménye, stb.) jellemzőkre is. Jelen munka itt kapcsolódik az indexálási problémához, és a felismerés jellege is multimédiássá válik.

A munka során az alábbi fontosabb feladatokat kívánjuk megoldani:

- Képi események történései és jellege (pl. biometria) leírása alapján a legfontosabbak automatikus kiválasztása
- Mozgások, alakzatok, hangesemények és átmenetek a képsorozaton: statisztikai struktúrák
- Sztereó összerendelés
- Az események strukturált leírása és értelmezése
- Jellemző események statisztikai modellezése
- Öntanuló osztályozási eljárások az adatbázisok alapján

## **Tervezett munkaszakaszok:**

**2005-ben:** Videó események jellemzőinek kiválasztása, erről adatbázis készítése, kulcs-momentumok illetve alaptörténések kiválasztása

- Eredmény: Adatbázis és leírása, publikációk

**2006-ban:** A csoportosítások, eseménysorok alapján statisztikai modellek felállítása; a besoroláshoz szükséges statisztikai paraméterek mérésének kidolgozása

- Eredmény: Adatbázis-modellek, legalább 2 publikáció

**2007-ben:** Multimédiás információk összekapcsolása, döntési/felismerési eljárások kidolgozása;

- Eredmény: Felismerő eljárások és tesztelésük, legalább 2 publikáció, köztük SCI folyóirat

## **A munka során elért eredmények:**

### **1. év:**

Előállítottunk a különböző film-jelenetekhez és kültéri tesztfelvételekhez tartozó videó adatbázist, annotációs bejegyzésekkel és vágási információkkal.

Egy ilyen tulajdonság volt a fókuszált (rendezett film kiemelt részletei) objektum a filmen. Ezek a részletek a fókuszterkép automatikus kiemelése útján elemezhetőek. Az előállított fókuszterkép egynézetű képekből és videókból alkalmas a képtér szereplői relatív helyzetének meghatározására az általunk kidolgozott dupla-iterációs vak-dekonvolúciós eljárással. Az eljárással a képek indexeléséhez vagy a videók eseményeinek leírásához kapunk új hatékony módszert.

További eredmény a "Hangspektrogramok jellegzetes struktúráinak képi elemzése", amely a videóesemények térbeli hanghatásainak elemzését segíti.

Mozgásalapú kameraregisztrációs és biometrikus (lépésmintázat) alapú képösszerendeléses eljárásokat dolgoztunk ki kültéri és beltéri videófelvételek háromdimenziós parametrikus elemzéséhez [8].

### **2. év:**

*Az 2006 év során a képi és hang események kiértékelésében, valamint az adatbázisba szervezett képek és videók fókuszterképének kinyerésében és az ez alapján történő keresésben, továbbá a több kamerát és mikrofont tartalmazó terek analízisében értünk el eredményeket. Markov alapú statisztikai összerendelések és dimenziócsökkentéssel is kombinált tulajdonságkiemelések is részei voltak a munkának.*

#### *1. Főszereplő kiemelése rendezett filmrészlet esetén*

Rendezett filmek esetén meghatározandó a főszereplő. Ennek során feltételezzük, hogy főszereplő az, akire/amire a (rendezett) film operatőre fókuszál. Egy képsík alapján a fókusz meghatározása eddig nem volt ismert. Erre adtunk megoldást [7]-ben, ahol bevezettünk egy új eljárást egy kép relatív fókuszterképének meghatározásához, amit az elmúlt évben jelentősen továbbfejlesztettünk [1]. Egy kép (nézet) alapján mélységében szegmentálni tudjuk a képi alakzatokat. A képek mélységi elrendezése alkalmas a kép-adatbázisok indexelésére, keresésére is. Az [1] megadott PR görbék egyértelműen mutatják eljárásunk újszerű és hatékony voltát. Egy elvileg is új hiba-mértéket vezettünk be (hiba-szög), ami a vak-dekonvolúciós iterációk során a hibának az ortogonalitási feltételtől való eltérését mutatja.

#### *2. Hosszú távú háttérmodell előállítása mozgó kamerán*

Nem rendezett, pl. biztonsági kamerafelvételek esetén viszont éppen az a cél, hogy megtaláljuk az eredetileg automatikusan szkennelő kamera mozgásában azt a pontot, amikor valaki beavatkozik, és kézzel kezdi mozgatni a kamerát. Ehhez a kameramozgás olyan elemzésére van szükség [10], ami eddig nem létezett a gyakorlatban. Amennyiben ez a beavatkozás mérhető, akkor már lehetőség van ennek az akciónak a bevonására a megerősített tanulásba.

Három problémát vizsgáltunk itt meg. Az első az, hogy hogyan tudjuk elkülöníteni egymástól a mozgó (mozgó autó, gyalogos stb.) és álló objektumokat (közlekedési táblák, házak, parkoló autó stb.). Álló kamera esetében ez viszonylag egyszerű, azonban mozgó kamera esetében az adott képen minden objektum elmozdul az előző képhez viszonyítva. A második probléma annak megállapítása, hogy mikor történt kezelői közbeavatkozás, azaz a kamera nem azt a képet mutatja, mint amikor a normál, megszokott mozgását végezné. A harmadik probléma pedig a kamera pozíciójának becslése, hogyan tudjuk megállapítani, hogy a teljes megfigyelt területből mit lát éppen a kamera. A kidolgozott eljárás [10] jól jelzi a kezelői beavatkozást, miközben a szkennelő kamera mozgását a biometrikai járásazonosítási eljárásokban megismert periodicitási vizsgálat egy módosításával oldottuk meg.

Ebben a munkában vizsgáltuk meg a nagydimenziós tulajdonságkiemelők szabadsági fokának lehetséges csökkentését is, PCA-jellegű kiemelésekkel.

### 3. *A hangtér vizsgálata*

A képi tartalom elemzése mellett a kísérő hang elemzése, szegmentálását és minősítése is szükséges. Ez sok esetben akkor is elárulja a váratlan eseménynek helyét és idejét, amikor a képen nem jelentős a változás. A hang elemzésére a spektrogram szegmentálása ad jó eszközt [11]. A megválaszolható kérdések egyrészt a hang jellegéhez, spektrális tulajdonságaihoz (beszéd, zene, zaj, keverék), másrészt a hangforrások számához, többszernés hang esetén a hangforrás(ok) térbeli elhelyezkedéséhez kapcsolódnak. A folyamatos hangszegmensek, valamint a nem folyamatos, de azonos forráshoz rendelhető hangszegmensek azonosítása támogatja a képtartalom értelmezését és képváltásokon áthúzódó esetekben lehetőséget ad független képtartalmak összekapcsolására is.

### 4. *Az optikai tér statisztikai vizsgálata*

Mint [3]-ban megmutattuk, a mozgás-gyanús pontok statisztikai összehasonlításával az eseménytér szerkezetének geometrikai adataira tehetünk becslést előfeltevés és modell nélkül. [9]-ben bemutatunk egy eljárást, ahol az egynézetű képen keressük meg a tükör, illetve az

árnyékot adó fényforrás vetítési pontját. Az eljárás során statisztikai korrelációtérképekre illesztett parametrikus hipotézismodelleket optimalizálunk a vetítési hiba segítségével. A kapott eredmények jól jellemzik a színhely eseményterének elrendezését.

### 3. év:

*A 2007-ben a mérések többsége nagy adatbázisokon történő alapos teszteléssel történt.*

1. Automatikus képillesztő eljárásokat fejlesztettünk ki, amelyek általános megoldást biztosítanak széles bázistávolságú sztereó-képek illesztésére tetszőleges körülmények esetén. Az algoritmusokat valós körülmények között teszteltük, és kidolgoztuk a tár- és számítás-igényes eljárás gyorsítását, hatékonyabbá tételét is [3].
2. A videókép mozgó alakzatainak detektáláshoz részletes kutatást folytattunk a színmodellek befolyásának megállapítására. Azt találtuk, hogy a felismerési eredmény lényegesen függ a színmetrikától is [2].
- A kapcsolódó projektekben multikamerás eseménydetektáló keretrendszert fejlesztettünk, amely számos saját fejlesztésű szűrő segítségével hatékonyabb térfigyelést tett lehetővé. Annotált videó-adatbázisokban történő visszakereső rendszert fejlesztettünk, amely szöveges kérdés és minta alapján tartalom alapú keresésekre képes. Ebben a keretrendszerben tudtuk tesztelni az elvileg új algoritmikus megoldásainkat is.
3. A képi tulajdonságok automatikus kinyeréséhez videó-feldolgozási szűrőket készítettünk, kép- és videó-indexelés és visszakeresés céljából, ill. szokatlan események automatikus detekciójának eléréséhez. Módszert dolgoztunk ki képek vizuális szempontból fontos területeinek automatikus elkülönítésére [1,2], és a módszer felhasználására képi visszakeresésben.  
Kifejlesztettünk egy módszert olyan automatikus nem-fotorealisztikus képalkotási eljárásra, amely a fontos területeket elkülönítve terület- és rétegfüggő festéssel készít grafikákat [4].
4. Kidolgoztunk egy módszert amely automatikusan szegmentál kis (néhány pixel területű) mozgó alakzatokat, track-elés és targeting céljára [5].
5. A mért tulajdonságokat, képi és video mozgásjellemzőket XML adatbázisba konvertáljuk, amit az általunk különböző célokra kifejlesztett kereső programok (annotációs és gépi indexelési paraméterek alapján) tudnak az adatbázishoz hozzárendelni.

## **A munkaterv az eredmények tükrében**

A projekt során a munkaterv főbb vonalait követtem, és az elért eredmények sokban megfelelnek az előzetes irányoknak. Néhány esetben azonban elkalandoztunk tudományosan érdekes, kihívást jelentő problémák kapcsolódó megoldásaihoz is. Olyan kérdésekben, mint: fókuszált főszereplő keresése, szinkérdések az előtér detekció tisztázása, és a sztochasztikus kölcsönhatások felfedezése esetén az elért eredmények túlmutatnak a korábbi elképzeléseken. Sikertült olyan, nemzetközileg is fontos eredményeket elérni, amiket korábban nem (tudtunk) tervezni. A tervezettnél több publikáció: **11** (ebből 6-ban OTKA hivatkozással) jelent meg a befejezésig, ebből 6 (3) SCI folyóirat, 15 (8) körüli impakt faktorral.

## **Személyi változások**

A munka kezdetén két pályázó név volt: én és Radványi András, valamint doktorandusz diákok. A PhD diákok közben doktoráltak, és már kutatóként vettek részt a munkában (Kovács Levente, Szilávik Zoltán), illetve Radványi András a rá jutó munkafázis elvégzése után már nem szerepelt a munkában. A munkában egyetemisták is részt vettek: annotációs és kódolási, tesztelési feladatokban.

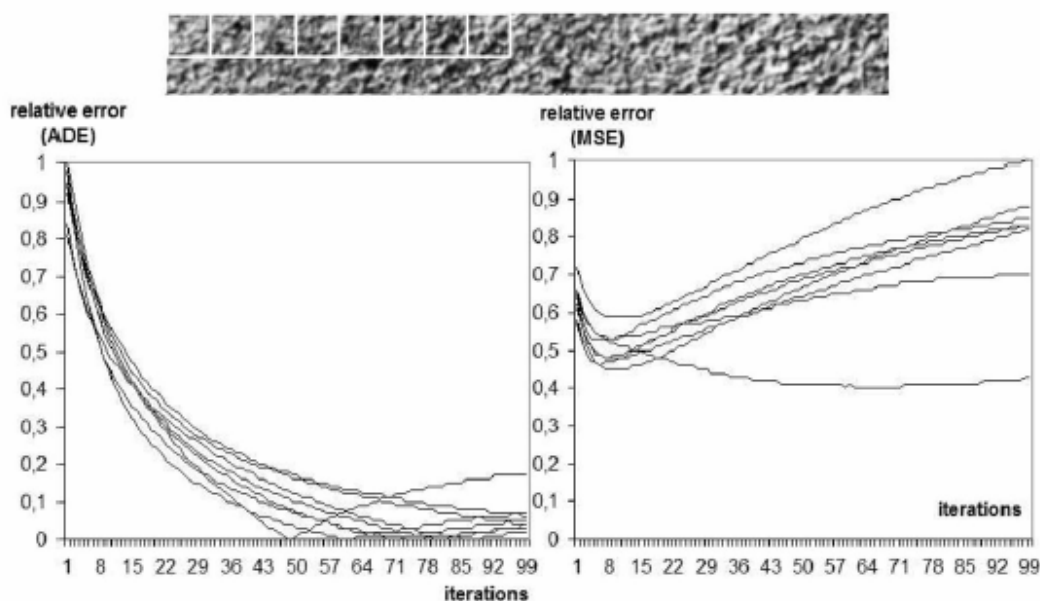
## Függelék: Néhány fontosabb és jellemző eredmény rövid bemutatása

A publikációk részletesen tartalmazzák a tudományos eredményeket. A jobb áttekinthetőség érdekében itt kiemeltem néhány jellemző részt, lehetőleg képekkel.

### Fókuszterkép mérése a főszereplő szegmentálásához

Bevezettünk egy vak dekonvolúción alapuló automatikus eljárást, amely képes képek relatív fókuszterképnek kinyerésére.

- Kísérletileg igazoltuk, hogy a dekonvolúciós fókuszterkép kinyerés különféle textúrált felületeket Megmutattuk a vak dekonvolúciós relatív fókuszterképek gyakorlati alkalmazhatóságát fókusz-alapú képi osztályozáson keresztül. Ezáltal lehetővé tettük képi adatbázisok indexelését fókuszterületek alapján.
- Megmutattuk, hogy a kinyert fókuszterképek szegmentációs képességét nagymértékben javítja, ha nem a négyzetes hibát, hanem egy általunk definiált ú módon, a dekonvolúciós eljárás maradékhibájának és a mintavett jelnek a merőlegességét (függetlenségét) mérjük. Ez *az elvileg is új* hibaszámítás kiküszöböli a négyzetes hiba és hasonló eljárások skálázási és nullponti problémáit.



Fent: 8 kiválasztott és elkent textúrált terület, amikre a vak deonvolúción alapuló fókuszterkép becslét számítottuk

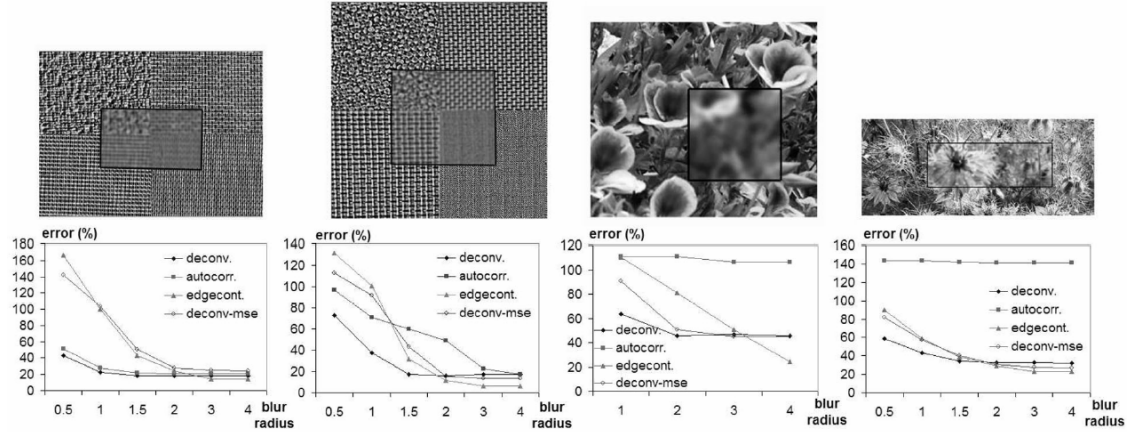
Balra: Az általunk definiált merőlegességi (függetlenségi) hiba értéke az iterációszám függvényében a 8 területre

Jobbra: Ugyanez négyzetes hibával számolva. Az irodalom is azt jegyzi, hogy 5-ös iteráció után divergálni kezd a négyzetes hiba. Neműgy, mit a baloldali új megoldásunknál.

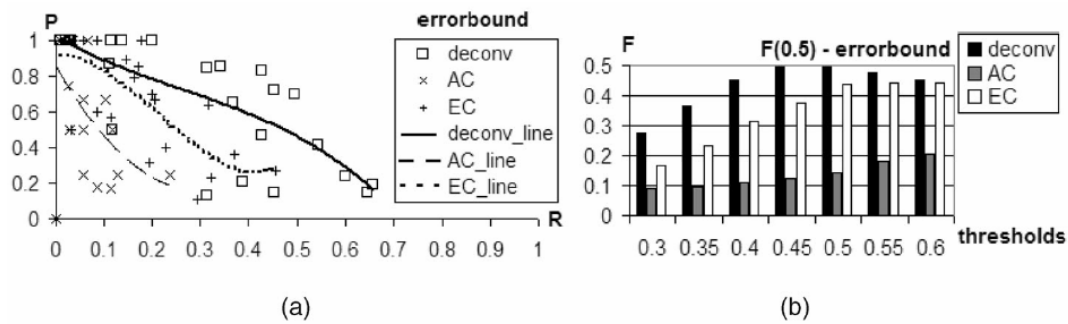


Az átalunk definiált merőlegességi (függetlenségi) hiba definíciója, ahol  $g$  a mért kép,  $g_k = h_k^*$  pedig a becslt  $f_k$  eredeti kép és a becslt  $h_k$  torzításból számolt becslt mérési eredmény:

$$E(g, g_k) = \left| \arcsin \frac{\langle g - g_k, g \rangle}{|g - g_k| \cdot |g|} \right|$$



Szegmentációs eredmények (hiba%, függőleges tengely) elmosódott (középen) textúrák elmosódási értékének (vízszintes tengely). „deconv” jelenti a mi eljárásunkat, míg a „deconv-mse” a négyzetes hiba alapút.



(a) Átlagos Precision-Recall görbe a küszöb feletti lekérdezés válaszaival (b) Az öt legjobb válasz alapján számolt F értékek (harmonikus Precision/Recall átlag)

## **Tükör és árnyék vetítési pontjának meghatározása statisztikus optimalizációval, előzetes modell nélkül**

*Megmutattuk, hogy a változás - mint információ - összegzése statisztikai értelemben jól modellezhető eloszlást eredményez, ami robusztus és felhasználható geometriai modellezési feladatokban. Eljárást adtunk a statisztikák parametrikus feldolgozására és felhasználására a kameranézetek különféle geometriai jellemzőinek meghatározásához [9].*

A környezet változásainak észlelése a kamera síkjában jellemzi a kamera környezethez viszonyított helyzetét és a környezet geometriai sajátosságait. Kültéri és beltéri kameraképek esetében egyaránt előforduló szituációk az árnyék, horizont és speciális esetben a tükröződő felületek meghatározása. Mindhárom eset definiálható geometriai modell optimalizációs feladatként melyek megoldásához szükséges mérési adathalmazt a mozgásstatisztikák parametrikus leíróiból nyertem. Összegezve, a vizsgált (együtt)mozgási statisztikák úgynevezett lokális értelemben számolva alkalmasak a kamera nézet analízisére, zajos, ismeretlen környezeti feltételek mellett.

## **Mozgásstatisztikák modell alapú feldolgozása robusztus és subpixel pontosságú jellemzőkinyerést tesz lehetővé.**

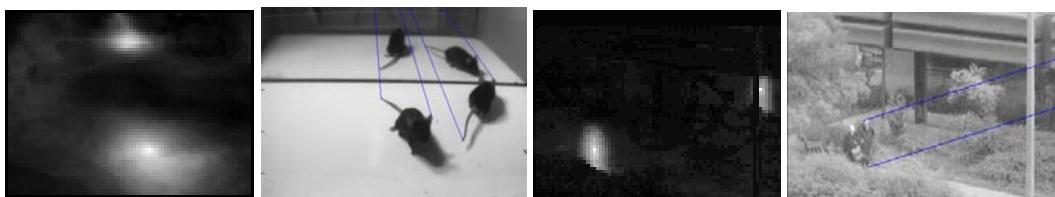
*Megmutattuk és kísérleti úton is igazoltuk, hogy a mozgásstatisztikák modell alapú leírása alkalmas a geometriai modellek meghatározásához szükséges pozíció-információ subpixel pontosságú és robusztus kinyerésére.*

A 2D-s (együtt)mozgási statisztikák az együttes események bekövetkezési gyakoriságát leíró feltételes valószínűségek közelítő becslései. Ezek modell alapú leírásának feltételét definiáltuk és a feltétel teljesülését különböző környezetekben megvizsgáltuk. Az elméleti eredmények alátámasztják a kísérleti tapasztalatokat, miszerint a módszer széles tartományban jól használható és a megkívánt pontosság eléréséhez szükséges bemeneti adatmennyiség jól becsülhető a környezeti (mozgásintenzitás) és algoritmikus (detekciós hiba) paraméterek alapján.

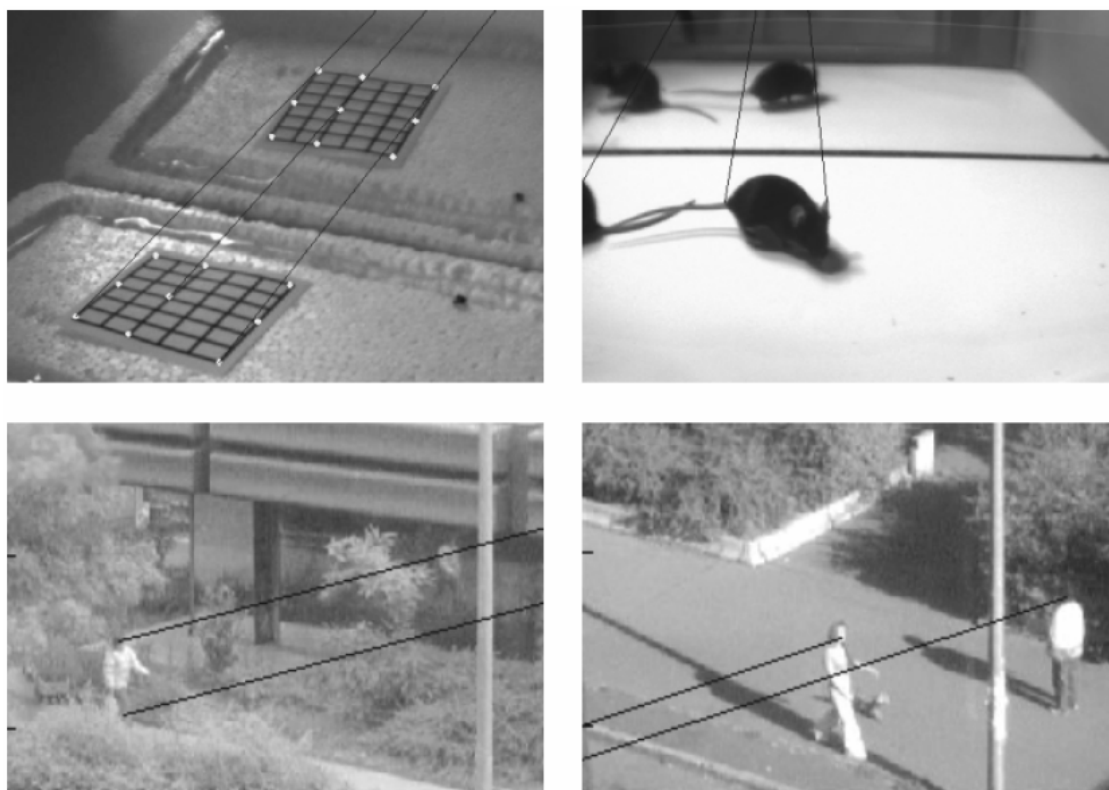
## **A sík, tükröződő felület és az árnyék geometriai modelljének meghatározása globális optimumkeresési módszerrel.**

*Megmutattuk, hogy a sík tükröződő felület és az árnyék geometriája megadható ferdén szimmetrikus fundamentális mátrix-al, ami egy pont-egyenes összerendelést jelent és két paraméterrel – a tükröközpont pozíciójával – egyértelműen megadható.*

Jósági függvényt definiáltunk a statisztikai információk felhasználásával. A függvény maximumának argumentuma adja a keresett modell paramétert, ami globális optimumkeresési feladathoz vezet. A módszer hatékonyságát és pontosságát kísérleti úton, kültéri és beltéri környezetben egyaránt igazoltuk. A bemeneti adatok térbeni felbontásának csökkentésével az eljárás futási sebessége jelentősen javítható, miközben az optimalizálás során meghatározott modell paraméterek számottevően nem változnak. A modell alapú feldolgozás ezen subpixeles pontosságot biztosító tulajdonságát kísérleti úton igazoltuk.



*Lokális együttmozgási statisztikák és az optimalizálással nyert model-paraméter (tükörközpont) demonstrálása bel- és kültéri környezetben.*



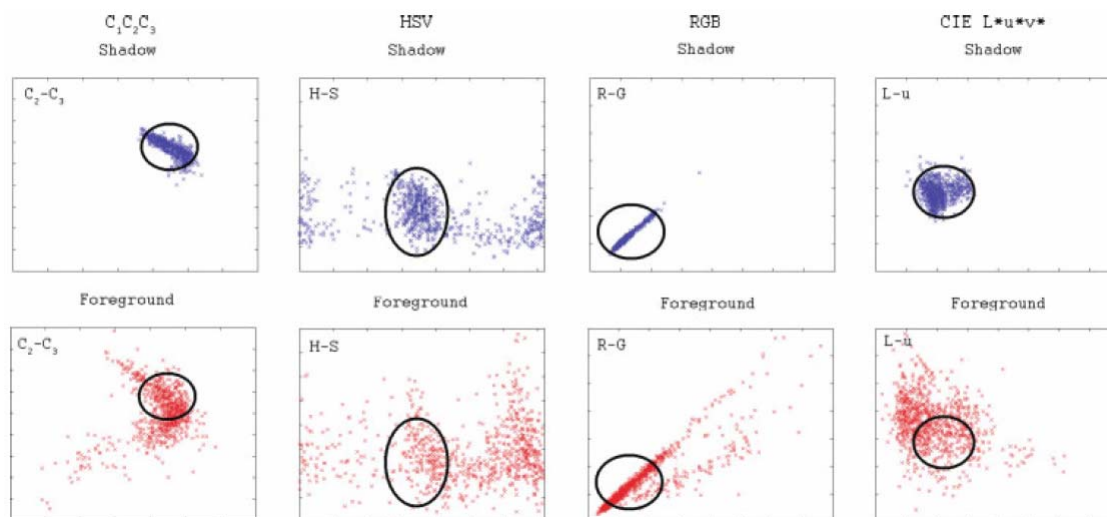
*Eredmények a mozgó alakzat és a tükörképe (illetve árnyéka) alapján számolt Vanishing Point (VP, vetítési pont) illetve vetítési egyenesek pontosságának igazolására*

## Színproblémák és metrikák

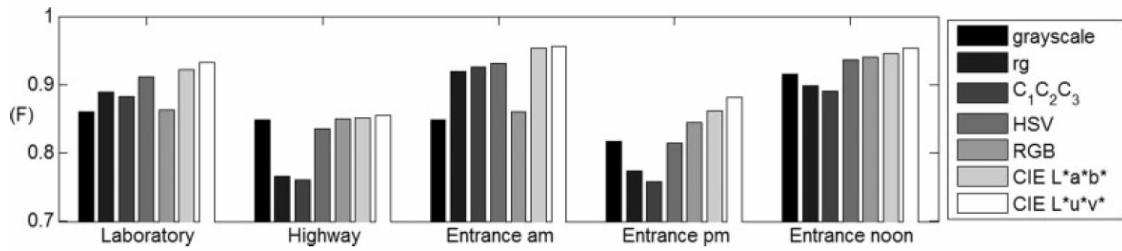
Kísérleti úton megmutattuk, hogy az elterjedt színterek közül a CIE  $L^*u^*v^*$  színmodell a leghatékonyabb az árnyék detekciójára mind a leíróterben végzett pixel szintű klaszterezés, mind a javasolt Markovi optimalizáció több színtérre történő általánosítása esetén.

Nyitott kérdés a szakirodalomban, hogy melyik színtérben érdemes az árnyékdetekciós vizsgálatokat elvégezni, és mekkora a helyes választás jelentossége. Megmutattuk, hogy olyan esetben, amikor a modellparaméterek számát praktikus okokból alacsonyan kell tartani, megfelelő színtér kiválasztásával a szegmentálás minősége lényegesen javítható.

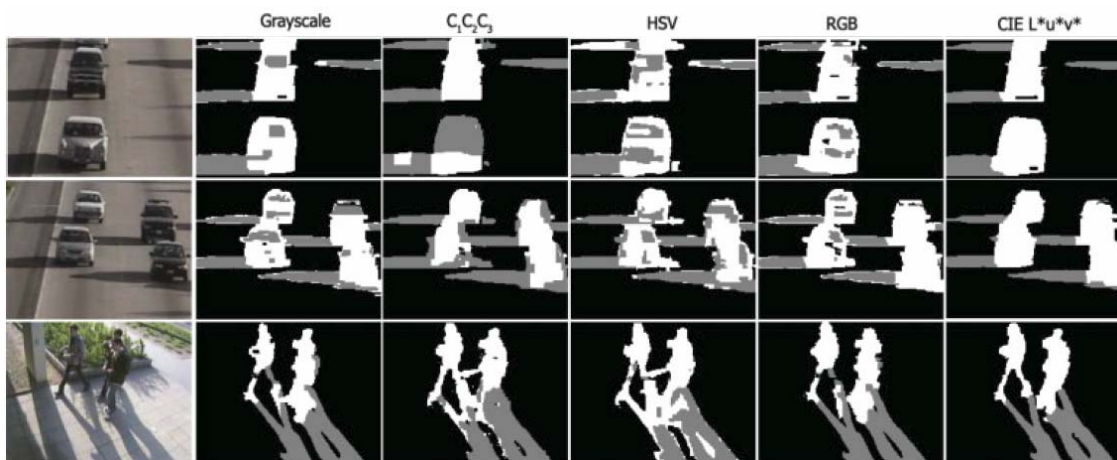
Terveztünk egy tisztán pixel szintű döntéseken alapuló, több színtérrel is együttműködni képes előtér-árnyék osztályozót, és annak gyakorlati relevanciáját a szakirodalom részletes áttekintésével és elemzésével támasztottuk alá. Az osztályozó model és a korábban is használt teszt adatbázis segítségével a színtereket kísérleti úton összehasonlítottuk. Mindkét kísérleti megközelítés a CIE  $L^*u^*v^*$  színtér egyértelmű dominanciáját hozta.



*Az egyes színterek szegmentációs képessége árnyék (fent) és előtér (lent) mintázatokra. A fekete ellipszis az optimalizált döntési paramétert mutatja.*



*A determinisztikus modell kiértékelése az  $F$  mérőszámmal (Precision/Recall harmonikus átlag) megadva a különböző videórészletekre*



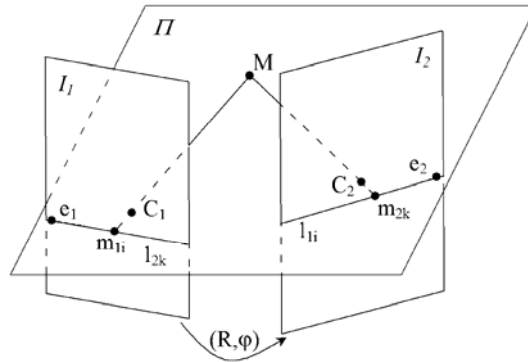
*MRF szegmentálsi eredmények a különböző színmodellekre*

Példaként mutatunk egy, a programrész által szolgáltatott tipikus XML formátumú kimenetet, egy adott videoklipp feldolgozása után:

```
<video_name> traffic6.avi </video_name>
<frame_width>320</frame_width>
<frame_height>240</frame_height>
<trackmask>traffic6_trackmask.bmp</trackmask>
<roimask>traffic6_roimask.bmp</roimask>
<shadow_darkness>0.730</shadow_darkness>
<shadow_dir_type>1</shadow_dir_type>
<shadow_vector_x>25</shadow_vector_x>
<shadow_vector_y>19</shadow_vector_y>
<din_area_mask>03_dynamicVideoPartMask.bmp</din_area_mask>
<static_shadow_mask>03_staticShadowMask.bmp</static_shadow_mask>
<din_area_ratio>0.709</din_area_ratio>
<detected_faces_ratio>0.000</detected_faces_ratio>
<detected_cars_ratio>0.886</detected_cars_ratio>
```

## ***Kamerák regisztrációja akkumulált mozgás-statisztikák segítségével***

Az együttmozgási statisztikákra alapozva új eljárást adtunk meg kamerák regisztrációjára.



Tekintsünk két kameraképet, átfedő nézetűvel. Legyen  $P(m_{ij})$  a valószínűsége annak, hogy az  $i$ . képnézet  $j$  pontja a közös területen van:  $m_{ij} \in OFV_i$ . Legyen  $P(m_{1i} | m_{2k})$  a feltételes valószínűsége annak, hogy az  $1$ . kép  $i$  pontja benne van a közös területben  $m_{1i} \in OFV_1$ , amikor a  $2$ . kép  $k$  pontjára ( $m_{2k}$  is), és a két pont ugyanarra a 3D pontra vonatkozik. Vagyis  $P(m_{1i} | m_{2k})$  jelenti annak esélyét, hogyha  $m_{2k}$ -ben mozgást tapasztalunk, akkor  $m_{1i}$ -ben is az van.

A  $t$  időbeli mozgás küszöbölt binaris értékét  $b(t)$ -vel jellemezve:

$$P(m_{1i} | m_{2k}) = \frac{1}{\sum_{t=1}^T \underline{b}_{2k}(t)} \sum_{t=1}^T \underline{b}_{1i}(t) \underline{b}_{2k}(t), \quad \sum_k P(m_{1i} | m_{2k}) = 1$$

Együttmozgási statisztikákat definiáltunk a különböző kameraképeken megfigyelt mozgások akkumulációjaként. Kísérletileg igazoltuk, hogy az együttmozgási statisztikák hatékonyan használhatók kamerák regisztrációjára emberi beavatkozás nélkül, mindenféle *a priori* környezeti információt figyelmen kívül hagyva. Az eljárás pontossága a *subpixel*-es tartományon belül van.

Megmutattuk, hogy az együttmozgási statisztikákon alapuló kameraregisztráció esetén a videók időben egymáshoz illeszthetők a regisztráció hibáját minimalizálva. A videók időbeni illesztésére kidolgozott eljárás szinkronizálatlan kamerák esetére a kamerák regisztrációjakor keletkező hibát méri, és ezt minimalizálva határozza meg az időbeni eltérést a videók között.

*Automatikus Bayes-i eljárás kamerák regisztrációjára tetszőleges struktúrájú objektumok, látványok esetére.*

Sokkamerás rendszerek által szolgáltatott képek illesztésének előfeltétele a kameraképek közös területeinek ismerete/észlelése. Új Bayes-i eljárást dolgoztunk ki kamerák közös területeinek automatikus meghatározására, ahol a modell megoldása ekvivalens egy periodikus Markov lánc határvalószínűségeinek meghatározásával. A kidolgozott eljárás alkalmas a kameraképek közös részeinek automatikus detekciójára mellőzve mindenféle emberi beavatkozást vagy *a priori* információt a megfigyelt objektumokról.

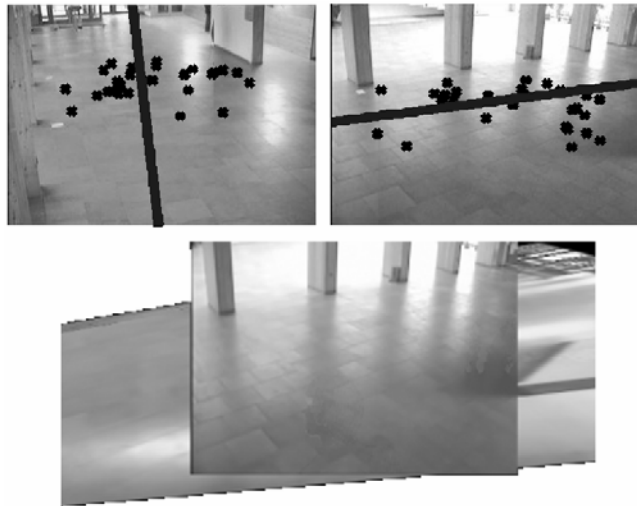


*Egyazon helyszínről készített képek egymásra illesztése.*

A szakirodalomból ismert képillesztő eljárások azt feltételezik, hogy a megfigyelt objektumok elég messze vannak a kamerától ahhoz, hogy síkon mozgó pontoknak tekinthessük őket. Ezáltal a feladat két sík összeillesztésére egyszerűsödik. Ez a feltételezés viszont csak kültéri rendszerek esetében érvényes. Beltéren a mozgó objektumok viszonylag közel vannak a kamerákhoz és a detektált mozgások a 3 dimenziós térben figyelhetők meg, azaz nem egy síkon. Viszont, ha olyan pontokat észlelünk melyek nagy valószínűséggel egy síkon (pl.: a padló síkja) mozognak, akkor továbbra is élhetünk a korábbi feltételezésünkkel.

Megmutattuk, hogy a mozgástörténetek entrópiájának vizsgálatával hatékony eljárás készíthető a jellemző mozgásokhoz tartozó képpontok kinyerésére. A mozgástörténetek entrópiáját vizsgáló eljárást dolgoztunk ki a megfigyelt mozgásokhoz tartozó képpontok

klasszifikációjára: (i) jelentős, nem folyamatos mozgások; (ii) folyamatos mozgások vagy kamerazaj. Kísérletileg megmutattuk, hogy egy árnyék-detekciós algoritmust az együttmozgások detekciójával kombinálva hatékony eljárás készíthető beltéri kamerák regisztrációjára.



*Árnyék-mozgás alapú képillesztés*

Az árnyékok ideális jellemzők beltéri kameraképek illesztéséhez; a legtöbb esetben a földön vagy padlón vannak, vagyis egy jól behatárolható síkon. Megmutattuk, hogy a mozgó objektumok árnyékának detekciójával hatékony eljárás készíthető beltéri kamerák regisztrációjára.

### ***Mozgó kamera képének szegmentációs problémái***

Kidolgoztuk a mozgó és álló objektumokat elkülönítését mozgó kamera esetére, miközben azt is becsülni kell, hogy a teljes megfigyelt területből mit lát éppen a kamera. A kidolgozott eljárás [10] jól jelzi a kezelői beavatkozást, miközben a szkennelő kamera mozgását a biometrikai járásazonosítási eljárásokban megismert periodicitási vizsgálat egy módosításával oldottuk meg.





*Felsőkép: mozaikkép gyenge minőségű webkamera használatakor. Alsó kép: Morfológiával eltüntettük a fényviszonyok változásából adódó zajokat, csak a mozgó objektum jelent meg*

## OTKA támogatás feltüntetésével készült publikációk:

1. L. Kovács , T. Szirányi, "Focus Area Extraction by Blind Deconvolution for Defining Regions of Interest", **IEEE Tr. Pattern Analysis and Machine Intelligence**, V.29, No.6, pp.1080-1085, **2007**
2. Cs. Benedek , T. Szirányi: "Study on Color Space Selection for Detecting Cast Shadows in Video Surveillanc", **International Journal of Imaging Systems and Technology**, Special Issue on Applied Color Image Processing. Vol. 17, No. 3. pp. 190-201, **2007**
3. Z. Szilávik, T. Szirányi, L. Havasi, "Stochastic view registration of overlapping cameras based on arbitrary motion", **IEEE Tr. Image Processing**, Vol.16, No.3, pp.710 - 720, **2007**
4. L. Kovács, T. Szirányi: "Painterly effects rendering with focus based level of detail", **Eurographics 2007**, Prague, **2007**, pp. 81-84.
5. L. Kovács, T. Szirányi: "Recognition of hidden pattern with background", Signal and data processing of small targets, **SPIE optics+photonics**, San Diego, **2007**. Vol.SPIE 6699, pp. 669906-1-8
6. Kovács L, Szirányi T: **2D Multilayer Painterly Rendering with Automatic Focus Extraction**, WSCG 2006, Full Papers Proceedings. Plzen, **2006**

## További kapcsolódó publikációk:

Megjegyzés: a 3 oldalas letters-ek és egyes konferenciakiadványok esetén a szűkös terjedelem miatt nem került be *egyáltalán* köszönet-szekció

7. L. Kovács, T. Szirányi: "Relative Focus Map Estimation Using Blind Deconvolution", **Optics Letters**, Vol.30, pp. 3021-3023, November **2005**
8. Cs. Benedek, L. Havasi, T. Sziranyi, Z. Szilavik: "Motion-Based Flexible Camera Registration", **IEEE International Conference on Advanced Video and Signal-Based Surveillance**, Como, IEEE, pp.439-444, **2005**
9. L. Havasi, T. Szirányi, "Estimation of Vanishing Point in Camera-Mirror Scenes Using Video", **Optics Letters**, Vol. 31, No. 10, pp: 1411-1413, **2006**
10. Losteiner Dávid, Karba Krisztián, Havasi László, Szirányi Tamás: "Mozgó kamerák képeinek feldolgozása", **KÉPAF 2007** Konferencia, Debrecen, 283-292. oldal
11. Zoltán Fodrócz, András Radványi: "Computational auditory scene analysis in cellular wave computing framework", **Int. J. on CTA**, p 489-515, Volume 34, Issue 4, July/August **2006**